Supplementary material: Kernelized Stein Discrepancy Tests of Goodness-of-Fit for Time-to-Event Data

A. Proofs and Derivations

A.1. Proofs of Section 3.1: Survival Stein Operator

A.1.1. Proof of Proposition 2

Let $\omega \in \mathcal{H}^{(s)}$. Then

$$\mathbb{E}_0((\mathcal{T}_0\omega)(T,\Delta) - (\mathcal{T}_0^s\omega)(T,\Delta)) = \mathbb{E}_0\left(\omega(T)\left[\Delta\left(\frac{f_0'(T)}{f_0(T)} - \lambda_C(x)\right) - \left(\Delta\frac{\lambda_0'(T)}{\lambda_0(T)} - \lambda_0(T)\right)\right]\right) \tag{23}$$

Observe that

$$\mathbb{E}_0\left(\Delta\omega(T)\lambda_C(T)\right) = \int_0^\infty \omega(x)\frac{f_C(x)}{S_C(x)}S_C(x)f_0(x)dx = \int_0^\infty \omega(x)\frac{f_0(x)}{S_0(x)}S_0(x)f_C(x)dx = \mathbb{E}_0((1-\Delta)\omega(T)\lambda_0(T)),$$

therefore, the RHS of Equation (23) is equal to

$$\mathbb{E}_0\left(\omega(T)\Delta\left(\frac{f_0'(T)}{f_0'(T)} + \lambda_0(T) - \frac{\lambda_0(T)}{\lambda_0(T)}\right)\right).$$

Finally, the last expectation is 0 due to the identity $\frac{f_0'(x)}{f_0(x)} = \frac{\lambda_0'(x)}{\lambda_0(x)} - \lambda_0(x)$, which follows from a simple computation.

A.1.2. Proof of Proposition 3

By definition,

$$\begin{aligned}
\sup_{\omega\in B_1(\mathcal{H})}\frac{1}{n}\sum_{i=1}^n(\mathcal{T}_0^{(s)}\omega)(T_i,\Delta_i) - (\mathcal{T}_0\omega)(T_i,\Delta_i) &= \sup_{\omega\in B_1(\mathcal{H})}\frac{1}{n}\sum_{i=1}^n\omega(T_i)\left(\Delta_i\lambda_C(T_i) - (1-\Delta_i)\lambda_0(T_i)\right)\\
&= \sup_{\omega\in B_1(\mathcal{H})}\left\langle\omega, \frac{1}{n}\sum_{i=1}^n K(T_i,\cdot)\left(\Delta_i\lambda_C(T_i) - (1-\Delta_i)\lambda_0(T_i)\right)\right\rangle_{\mathcal{H}}\\
&= \left\|\frac{1}{n}\sum_{i=1}^n K(T_i,\cdot)\left(\Delta_i\lambda_C(T_i) - (1-\Delta_i)\lambda_0(T_i)\right)\right\|_{\mathcal{H}}
\end{aligned}$$

We continue by proving that the previous norm converges to zero in probability. Observe that by the symmetrization lemma (**?**)Lemma 6.4.2]vershynin2019high, it holds

$$\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^n K(T_i,\cdot)\left(\Delta_i\lambda_C(T_i) - (1-\Delta_i)\lambda_0(T_i)\right)\right\|_{\mathcal{H}}\right] \leq 2\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^n W_iK(T_i,\cdot)\left(\Delta_i\lambda_C(T_i) - (1-\Delta_i)\lambda_0(T_i)\right)\right\|_{\mathcal{H}}\right]$$

where $W_1,\ldots,W_n$ are i.i.d. Rademacher random variables, independent of the data $(T_i,\Delta_i)_{i=1}^n$. Then, by Jensen's inequality, and by using that $\mathbb{E}(W_i)=0$, we conclude that the previous expression converges to zero in probability, as

$$\mathbb{E}\left[\left\|\frac{1}{n}\sum_{i=1}^n W_iK(T_i,\cdot)\left(\Delta_i\lambda_C(T_i) - (1-\Delta_i)\lambda_0(T_i)\right)\right\|_{\mathcal{H}}^2\right] = \mathbb{E}\left[\frac{1}{n^2}\sum_{i=1}^n K(T_i,T_i)\left(\Delta_i\lambda_C(T_i) - (1-\Delta_i)\lambda_0(T_i)\right)^2\right] \to 0,$$

a.s., where the limit result holds due the law of large numbers which can be applied under the Condition in Equation (12) and since $|K(x,y)| \leq c_1$, as

$$\begin{aligned}
\mathbb{E}\left[K(T_i,T_i)\left(\Delta_i\lambda_C(T_i) - (1-\Delta_i)\lambda_0(T_i)\right)^2\right] &\leq c_1\mathbb{E}\left[\left(\Delta_i\lambda_C(T_i)^2 + (1-\Delta_i)\lambda_0(T_i)^2\right)\right]\\
&= c_1\int_0^\infty\left(\lambda_C(x) + \lambda_0(x)\right)f_C(x)f_0(x)dx < \infty.
\end{aligned}$$

A.2. Proofs of Section 3.3: Proportional Stein Operator

PROOF OF PROPOSITON 4

We start by claiming that the following equation holds true for every $\omega \in \mathcal{H}^{(s)}$:

$$\frac{1}{n}\sum_{i=1}^{n}\left((\widehat{\mathcal{T}}_0^{(p)}\omega)(T_i,\Delta_i)-(\mathcal{T}_0^{(p)}\omega)(T_i,\Delta_i)\right)\xrightarrow{\mathbb{P}}0. \tag{24}$$

Then, the main result follows from Equation (24), by using the law of large numbers and that

$$\mathbb{E}_0\left[(\mathcal{T}_0^{(p)}\omega)(T_1,\Delta_1)\right]=\int_0^\infty\frac{(\omega(t)\lambda_0(t))'}{\lambda_0(t)}\frac{1}{S_0(t)S_C(t)}S_C(t)f_0(t)dt=\int_0^\infty\frac{(\omega(t)\lambda_0(t))'}{\lambda_0(t)}\lambda_0(t)dt=0,$$

which follows from the definition of our operator (see Equation (18)).

We finish the proof by proving our claim in Equation (24). Observe that

$$\left|\frac{1}{n}\sum_{i=1}^{n}\left((\widehat{\mathcal{T}}_0^{(p)}\omega)(T_i,\Delta_i)-(\mathcal{T}_0^{(p)}\omega)(T_i,\Delta_i)\right)\right|\leq\frac{1}{n}\sum_{i=1}^{n}\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\left|\frac{\Delta_i}{Y(T_i)/n}-\frac{\Delta_i}{S_T(T_i)}\right|, \tag{25}$$

where $S_T(t)=S_C(t)S_0(t)$ holds under the null hypothesis. We proceed to prove that the previous sum tends to 0 in probability when n grows to infinity. Let $\varepsilon>0$ and define $t_\varepsilon>0$ as the infimum of all t such that $\int_t^\infty|(\omega(x)\lambda_0(x))'|\,dx<\varepsilon$. Notice that such t_ε is well-defined since $\int_0^\infty|(\omega(x)\lambda_0(x))'|\,dx<\infty$. We continue by splitting the sum in Equation (25) into two regions, $\{T_i\leq t_\epsilon\}$ and $\{T_i>t_\epsilon\}$, obtaining that Equation (25) equals

$$\frac{1}{n}\sum_{i=1}^{n}\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\left|\frac{\Delta_i}{Y(T_i)/n}-\frac{\Delta_i}{S_T(T_i)}\right|\mathbb{1}_{\{T_i\leq t_\varepsilon\}}+\frac{1}{n}\sum_{i=1}^{n}\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\left|\frac{\Delta_i}{Y(T_i)/n}-\frac{\Delta_i}{S_T(T_i)}\right|\mathbb{1}_{\{T_i>t_\varepsilon\}}, \tag{26}$$

and we prove that both sums tend to 0 in probability when n grows to infinity. We start with the first term. Observe that

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\left|\frac{\Delta_i}{Y(T_i)/n}-\frac{\Delta_i}{S_T(T_i)}\right|\mathbb{1}_{\{T_i\leq t_\varepsilon\}}&\leq\sup_{t\leq t_\epsilon}\left|\frac{1}{Y(t)/n}-\frac{1}{S_T(t)}\right|\frac{1}{n}\sum_{i=1}^{n}\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\Delta_i\mathbb{1}_{\{T_i\leq t_\varepsilon\}}\\&=o_p(1),\end{aligned}$$

where the previous result holds since $\sup_{t\leq t_\epsilon}\left|\frac{1}{Y(t)/n}-\frac{1}{S_T(t)}\right|\to 0$ almost surely by the Glivenko-Cantelli Theorem, and since

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\Delta_i 1_{\{T_i\leq t_\epsilon\}}\to\mathbb{E}\left[\frac{|(\omega(T_1)\lambda_0(T_1))'|}{\lambda_0(T_1)}\Delta_1\mathbb{1}_{\{T_1\leq t_\varepsilon\}}\right]&=\int_0^{t_\epsilon}\frac{|(\omega(t)\lambda_0(t))'|}{\lambda_0(t)}S_C(t)f_0(t)dt\\&=\int_0^{t_\epsilon}|(\omega(t)\lambda_0(t))'|\,dt<\infty,\end{aligned}$$

where the last expression is finite due to Equation (17).

Next, we deal with the second term in equation (26). Theorem 3.2.1. of Gill (1980) yields $\sup_{t\leq\tau_n}\left|1-\frac{Y(T_i)/n}{S_T(T_i)}\right|=O_p(1)$, where $\tau_n=\max\{T_1,\ldots,T_n\}$, and, Lemma 2.7 of Gill (1983) yields $\sup_{t\leq\tau_n}nS_T(t)/Y(t)=O_p(1)$ (recall that $S_T(t)=S_0(t)S_C(t)$). From the previous results, we get

$$\begin{aligned}\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\left|\frac{1}{Y(T_i)/n}-\frac{1}{S_T(T_i)}\right|\mathbb{1}_{\{T_i>t_\varepsilon\}}&=\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\frac{1}{Y(T_i)/n}\left|1-\frac{Y(T_i)/n}{S_T(T_i)}\right|\mathbb{1}_{\{T_i>t_\varepsilon\}}\\&=O_p(1)\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\frac{1}{Y(T_i)/n}\mathbb{1}_{\{T_i>t_\varepsilon\}}\\&=O_p(1)\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\frac{1}{S_0(T_i)S_C(T_i)}\mathbb{1}_{\{T_i>t_\varepsilon\}}.\end{aligned}$$

Now, notice that

$$\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{|(\omega(T_i)\lambda_0(T_i))'|}{\lambda_0(T_i)}\frac{1}{S_0(T_i)S_C(T_i)}\mathbb{1}_{\{T_i>t_\varepsilon\}}\overset{a.s.}{\to}\mathbb{E}_0\left[\Delta_1\frac{|(\omega(T_1)\lambda_0(T_1))'|}{\lambda_0(T_1)}\frac{1}{S_0(T_1)S_C(T_1)}\mathbb{1}_{\{T_1>t_\varepsilon\}}\right]$$
$$=\int_{t_\varepsilon}^{\infty}\frac{|(\omega(x)\lambda_0(x))'|}{\lambda_0(x)}\frac{f_0(x)S_C(x)}{S_0(x)S_C(x)}dx$$
$$=\int_{t_\varepsilon}^{\infty}|(\omega(x)\lambda_0(x))'|dx<\varepsilon,$$

where the first equality holds by Equation (4), and the last inequality comes from the definition of t_ε. Since we can choose $\varepsilon>0$ as small as desired, we conclude the result.

A.3. Proofs Section 4: Censored-Data Kernel Stein Discrepancy

A.3.1. Proof of Proposition 5

Proof. Notice that, by the definition of the random function $\xi^{(c)}(\Delta,T)$, we have that $(\mathcal{T}^{(c)}\omega)(T,\Delta)=\langle\omega,\xi^{(c)}(T,\Delta)\rangle_{\mathcal{H}^{(c)}}$. Also notice that, $\xi^{(c)}(x,\delta)\in\mathcal{H}^{(c)}$ for each fixed (x,δ), and that the expectation, $\mathbb{E}_X\left[\xi^{(c)}(T,\Delta)\right]\in\mathcal{H}^{(c)}$ if and only if equation (21) is satisfied (the previous expectation has to be understood in the Bochner sense, as we are taking expectation of a random function).

Then,

$$\begin{aligned}\text{c-KSD}(f_X\|f_0)^2=\sup_{\omega\in B_1(\mathcal{H}^{(c)})}\mathbb{E}_X\left[(\mathcal{T}_0^{(c)}\omega)(T,\Delta)\right]^2&=\sup_{\omega\in B_1(\mathcal{H}^{(c)})}\mathbb{E}_X\left[\left\langle\omega,\xi^{(c)}(T,\Delta)\right\rangle_{\mathcal{H}^{(c)}}\right]^2\\&=\sup_{\omega\in B_1(\mathcal{H}^{(c)})}\left\langle\omega,\mathbb{E}_X\left[\xi^{(c)}(T,\Delta)\right]\right\rangle_{\mathcal{H}^{(c)}}^2\\&=\left\|\mathbb{E}_X\left[\xi^{(c)}(T,\Delta)\right]\right\|_{\mathcal{H}^{(c)}}^2\\&=\left\langle\mathbb{E}_X\left[\xi^{(c)}(T,\Delta)\right],\mathbb{E}_X\left[\xi^{(c)}(T',\Delta')\right]\right\rangle_{\mathcal{H}^{(c)}}\\&=\mathbb{E}_X\left[\left\langle\xi^{(c)}(T,\Delta),\xi^{(c)}(T',\Delta')\right\rangle_{\mathcal{H}^{(c)}}\right]\\&=\mathbb{E}_X\left[h^{(c)}((T,\Delta),(T',\Delta'))\right],\end{aligned}$$

where the third equality is due to the linearity of expectation and the inner product, the fourth equality follows from the definition of norm (and since we are taking supremum in the unit ball), and the second to last equality is, again, due to the linearity of the expectation and inner product. □

A.3.2. Explicit computation of $h^{(c)}$

Denote $\phi(x,\delta)=\delta\frac{\lambda_0'(x)}{\lambda_0(x)}-\lambda_0(x)$, and $L_1(x,y)=\frac{\partial}{\partial x}K^{(c)}(x,y)$, $L_2(x,y)=\frac{\partial}{\partial y}K^{(c)}(x,y)$ and $L=\frac{\partial^2}{\partial x\partial y}K^{(c)}(x,y)$. For simplicity of exposition, we will drop the superindex (c) in all cases.

Survival Stein operator $(c=s)$: For this case, we have

$$\begin{aligned}\xi(x,\delta)=(\mathcal{T}_0K)((x,\delta),\cdot)&=\delta\frac{\partial}{\partial x}K(x,\cdot)+\left(\delta\frac{\lambda_0'(x)}{\lambda_0(x)}-\lambda_0(x)\right)K(x,\cdot)+\lambda_0(0)K(0,\cdot)\\&=\delta L_1(x,\cdot)+\phi(x,\delta)K(x,\cdot)+\lambda_0(0)K(0,\cdot).\end{aligned}$$

Notice that a simple computation shows that $L(x,y)=\langle L_1(x,\cdot),L_1(y,\cdot)\rangle_{\mathcal{H}}$, then

$$\begin{aligned}h^{(s)}((x,\delta),(x',\delta'))=&\,\delta\delta'L(x,x')+\delta\phi(x',\delta')L_1(x,x')+\delta\lambda_0(0)L_1(x,0)\\&+\phi(x,\delta)\delta'L_2(x,x')+\phi(x,\delta)\phi(x',\delta')K(x,x')+\phi(x,\delta)\lambda_0(0)K(x,0)\\&+\lambda_0(0)\delta'L_2(0,x')+\lambda_0(0)\phi(x',\delta')K(0,x')+\lambda_0(0)^2K(0,0).\end{aligned}$$

Martingale Stein operator $(c = m)$: Observe that in this case

$$\xi(x,\delta) = (\mathcal{T}_0 K)((s,\delta),\cdot) = \frac{\delta}{\lambda_0(x)} L_1(x,\cdot) - K(x,\cdot) + K(0,\cdot).$$

Then, by the reproducing kernel property

$$\begin{aligned} h^{(m)}(x,\delta),(x',\delta')) = & \frac{\delta}{\lambda_0(x)}\frac{\delta'}{\lambda_0(x')} L(x,x') - \frac{\delta}{\lambda_0(x)} L_1(x,x') + \frac{\delta}{\lambda_0(x)} L_1(x,0) \\ & - \frac{\delta'}{\lambda_0(x')} L_2(x,x') + K(x,x') - K(x,0) \\ & + \frac{\delta'}{\lambda_0(x')} L_2(0,x') - K(0,x') + K(0,0). \end{aligned}$$

Proportional Stein operator $(c = p)$: Notice that, in this case, we use $\widehat{\mathcal{T}}_0^{(p)}$, given in Equation (16), to compute $\widehat{\xi}^{(p)}(x,\delta) = (\widehat{\mathcal{T}}_0^{(p)} K^{(p)})((x,\delta),\cdot)$ since $\mathcal{T}_0^{(p)}$ is not available, as it depends on S_C, which is unknown even under the null hypothesis. Then,

$$\widehat{\xi}(x,\delta) = (\widehat{\mathcal{T}}_0 K)((x,\delta),\cdot) = \left(L_1(x,\cdot) + \frac{\lambda_0'(x)}{\lambda_0(x)} K(x,\cdot)\right)\frac{\delta}{Y(x)/n}.$$

Define $K^\star(x,y) = \left(\frac{\partial^2}{\partial x \partial y}\lambda_0(x)\lambda_0(y)K(x,y)\right)$. Then, by the reproducing kernel property,

$$\widehat{h}^{(p)}((x,\delta),(x',\delta')) = n^2 \frac{\delta\delta'}{Y(x)Y(x')} K^\star(x,x').$$

Recall that $Y(t) = \sum_{k=1}^n \mathbb{1}_{\{T_k \geq t\}}$ denotes the risk function, which depends on all the data points, hence we write $\widehat{h}^{(p)}$ to recall the reader that this kernel is a random one.

A.4. Proofs of Section 5: Goodness-of-fit via c-KSD

The following lemmas show that, under Conditions c) and d) (depending on which case), the kernels $h^{(c)}$ have finite first and second moment. These moment conditions on the kernel are important to deduce asymptotic results.

Lemma 9. *Let (T',Δ') and (T,Δ) be independent samples from μ_X, and assume that Condition $d)$ holds. Then,*

$$\mathbb{E}_X\left[|h^{(c)}((T,\Delta),(T,\Delta))|\right] < \infty, \quad \textit{and} \quad \mathbb{E}_X\left[|h^{(c)}((T,\Delta),(T',\Delta'))|\right] < \infty$$

for $c \in \{s,m,p\}$, under the alternative hypothesis.

Lemma 10. *Let (T',Δ') and (T,Δ) be independent samples from μ_0, and assume that Condition $c)$ holds. Then*

$$\mathbb{E}_0\left[|h^{(c)}((T,\Delta),(T,\Delta))|\right] < \infty, \quad \textit{and} \quad \mathbb{E}_0\left[h^{(c)}((T,\Delta),(T',\Delta'))^2\right] < \infty$$

for $c \in \{s,m,p\}$, under the null hypothesis.

We just proof Lemma 9 since the proof of Lemma 10 is essentially the same.

Proof of Lemma 9. First of all, note that for any kernel (positive-definite function), it holds

$$h^{(c)}((x,\delta),(x',\delta')) \leq \frac{1}{2}h^{(c)}((x,\delta),(x,\delta)) + \frac{1}{2}h^{(c)}((x',\delta'),(x',\delta')),$$

hence, it is enough to only prove the first part of the lemma.

Survival Stein operator $(c = s)$: Recall $\xi^{(s)}(x,\delta) = \delta L_1(x,\cdot) + \phi(x,\delta)K(x,\cdot) + \lambda_0(0)K(0,\cdot)$, where $L_1(x,y) = \frac{\partial}{\partial x}K(x,y)$ $\phi(x,\delta) = \delta\frac{\lambda_0'(x)}{\lambda_0(x)} - \lambda_0(x)$, then

$$\begin{aligned}\mathbb{E}_X\left[\left|h^{(s)}((T,\Delta),(T,\Delta))\right|\right] &= \mathbb{E}_X\left[\left\|\xi^{(s)}(T,\Delta)\right\|^2_{\mathcal{H}^{(s)}}\right]\\ &\le 4\mathbb{E}_X\left[\|\Delta L_1(T,\cdot)\|^2_{\mathcal{H}^{(s)}} + \|\phi(T,\Delta)K(T,\cdot)\|^2_{\mathcal{H}^{(s)}}\right] + 4\,\|\lambda_0(0)K(0,\cdot)\|^2_{\mathcal{H}^{(s)}}\\ &\le 4\mathbb{E}_X\left[\|\Delta L_1(T,\cdot)\|^2_{\mathcal{H}^{(s)}}\right] + 4\mathbb{E}_X\left[\|\phi(T,\Delta)K(T,\cdot)\|^2_{\mathcal{H}^{(s)}}\right] + 4\lambda_0(0)^2K(0,0).\end{aligned}$$

The first and third term in the previous equation are finite under the technical Conditions a) and b). Thus, we only need to check

$$\mathbb{E}_X\left[\|\phi(T,\Delta)K(T,\cdot)\|^2_{\mathcal{H}^{(s)}}\right] = \mathbb{E}_X\left[\phi(T,\Delta)^2|K(T,T)|\right] < \infty,$$

which is guaranteed by Condition d).

Martingale Stein operator $(c = m)$: Recall that $\xi^{(m)}(x,\delta) = \phi(x,\delta)L_1(x,\cdot) - K(x,\cdot) + K(0,\cdot)$, where $L_1(x,y) = \frac{\partial}{\partial x}K(x,y)$ and $\phi(x,\delta) = \frac{\delta}{\lambda_0(x)}$. Then

$$\begin{aligned}\mathbb{E}_X\left[\left|h^{(m)}((T,\Delta),(T,\Delta))\right|\right] &= \mathbb{E}_X\left[\left\|\xi^{(m)}(T,\Delta)\right\|^2_{\mathcal{H}^{(m)}}\right]\\ &\le 4\mathbb{E}_X\left[\|\phi(T,\Delta)L_1(T,\cdot)\|^2_{\mathcal{H}^{(s)}}\right] + 4\mathbb{E}\left[\|K(T,\cdot)\|^2_{\mathcal{H}^{(s)}}\right] + 4\,\|K(0,\cdot)\|^2_{\mathcal{H}^{(s)}}.\end{aligned}$$

Observe that the second and third term are finite under Condition a). Additionally, define $L(x,y) = \frac{\partial^2}{\partial x \partial y}K(x,y)$ and notice that

$$\mathbb{E}_X\left[\|\phi(T,\Delta)L_1(T,\cdot)\|^2_{\mathcal{H}^{(s)}}\right] = \mathbb{E}_X\left[\phi(T,\Delta)^2L(T,T)\right] = \mathbb{E}_X\left[\frac{\Delta}{\lambda_0(T)^2}L(T,T)\right] < \infty$$

holds under Condition d) (Notice that $L = K^\star$ in Condition d.2)).

Proportional Stein operator $(c = p)$. This case follows directly from Condition d.3). □

A.4.1. PROOF OF THEOREM 6

We distinguish between two cases: first, when $h^{(c)}$ is a deterministic kernel (that is $c \in \{s, m\}$), and second, when $\widehat{h}^{(c)}$ is a random kernel, meaning $c = p$.

Deterministic kernel $(c \in \{s, m\})$**:** For the first case, we have

$$\widehat{\text{c-KSD}}^2(f_X||f_0) = \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n h^{(c)}((T_i,\Delta_i),(T_j,\Delta_j)),$$

which is a V-statistic of order 2. Thus, by using the law of large numbers for V-statistics, we deduce

$$\widehat{\text{c-KSD}}^2(f_X||f_0) \overset{a.s.}{\to} \mathbb{E}_X\left(h^{(c)}((T,\Delta),(T',\Delta'))\right) = \text{c-KSD}^2(f_X||f_0),$$

as n grows to infinity. Notice that the previous limit result requires the following conditions: $\mathbb{E}_X\left(|h^{(c)}((T,\Delta),(T,\Delta))|\right) < \infty$ and $\mathbb{E}_X\left(|h^{(c)}((T,\Delta),(T',\Delta'))|\right) < \infty$, which are satisfied under Condition d) by Lemma 9.

Random kernel $(c = p)$**:** For the second case, recall that

$$\widehat{\text{p-KSD}}^2(f_X||f_0) = \sum_{i=1}^n\sum_{j=1}^n \widehat{h}^{(p)}((T_i,\Delta_i),(T_j,\Delta_j)), \tag{27}$$

where $\widehat{h}^{(p)}$ is a random kernel. Our first step will be to assume that we can replace the random kernel $\widehat{h}^{(p)}$, given by $\widehat{h}^{(p)}((x,\delta),(x',\delta')) = n^2\frac{\delta\delta' K^\star(x,x')}{Y(x)Y(x')}$, by its limit $h^{(p)}((x,\delta),(x',\delta')) = \frac{\delta\delta' K^\star(x,x')}{S_T(x)S_T(x')}$, where $K^\star(x,y) = \left(\frac{\partial^2}{\partial x\partial y}K(x,y)\lambda_0(x)\lambda_0(y)\right)$. We claim that

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \widehat{h}^{(p)}((T_i,\Delta_i),(T_j,\Delta_j)) = \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n h^{(p)}((T_i,\Delta_i),(T_j,\Delta_j)) + o_p(1), \tag{28}$$

and then we have that

$$\begin{aligned}\widehat{\text{p-KSD}}^2(f_X\|f_0) &= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \widehat{h}^{(p)}((T_i,\Delta_i),(T_j,\Delta_j))\\ &= \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n h^{(p)}((T_i,\Delta_i),(T_j,\Delta_j)) + o_p(1)\\ &= \mathbb{E}_X(h^{(p)}((T,\Delta),(T',\Delta'))) + o_p(1) = \text{p-KSD}^2(f_X\|f_0) + o_p(1),\end{aligned}$$

where the third equality is due to the standard law of large numbers for V statistics, and by Condition d.3) and Lemma 9.

We finish the proof by proving the claim made in Equation (28). Recall that

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \widehat{h}^{(p)}((T_i,\Delta_i),(T_j,\Delta_j)) = \left\|\frac{1}{n}\sum_{i=1}^n \widehat{\xi}^{(p)}(T_i,\Delta_i)\right\|^2_{\mathcal{H}^{(p)}},$$

and

$$\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n h^{(p)}((T_i,\Delta_i),(T_j,\Delta_j)) = \left\|\frac{1}{n}\sum_{i=1}^n \xi^{(p)}(T_i,\Delta_i)\right\|^2_{\mathcal{H}^{(p)}}, \tag{29}$$

where $\widehat{\xi}^{(p)}(x,\delta) = n\frac{(K(x,\cdot)\lambda_0(x))'}{\lambda_0(x)}\frac{\delta}{Y(x)}$ and $\xi^{(p)}(x,\delta) = \frac{(K(x,\cdot)\lambda_0(x))'}{\lambda_0(x)}\frac{\delta}{S_T(x)}$. Then, by the triangular inequality, and by taking square (notice that $\|b\| - \|a-b\| \leq \|a\| \leq \|b\| + \|a-b\|$), the claim in Equation (28) follows from proving:

i) $\left\|\frac{1}{n}\sum_{i=1}^n \widehat{\xi}^{(p)}(T_i,\Delta_i) - \xi^{(p)}(T_i,\Delta_i)\right\|_{\mathcal{H}^{(p)}} = o_p(1)$, and

ii) $\left\|\frac{1}{n}\sum_{i=1}^n \xi^{(p)}(T_i,\Delta_i)\right\|_{\mathcal{H}^{(p)}} = O_p(1)$.

Notice that item ii) holds trivially by Equation (29), and by the law of large numbers for V-statistics, which can be applied due to Lemma 9, under Condition d). We finish by proving the result in item i). Following the same steps used in Equation (25), we have that

$$\begin{aligned}\left\|\frac{1}{n}\sum_{i=1}^n \widehat{\xi}^{(p)}(T_i,\Delta_i) - \xi^{(p)}(T_i,\Delta_i)\right\|_{\mathcal{H}^{(p)}} &= \left\|\frac{1}{n}\sum_{i=1}^n \frac{(K(T_i,\cdot)\lambda_0(T_i))'}{\lambda_0(T_i)}\left(\frac{\Delta_i}{Y(T_i)/n} - \frac{\Delta_i}{S_T(T_i)}\right)\right\|_{\mathcal{H}^{(p)}}\\ &= \sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{n}\sum_{i=1}^n \frac{(\omega(T_i)\lambda_0(T_i))'}{\lambda_0(T_i)}\left(\frac{\Delta_i}{Y(T_i)/n} - \frac{\Delta_i}{S_T(T_i)}\right)\\ &\leq \sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{n}\sum_{i=1}^n \frac{(\omega(T_i)\lambda_0(T_i))'}{\lambda_0(T_i)}\left(\frac{\Delta_i}{Y(T_i)/n} - \frac{\Delta_i}{S_T(T_i)}\right)\mathbb{1}_{\{T_i\leq t_\varepsilon\}} \quad (30)\\ &\quad+ \sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{n}\sum_{i=1}^n \frac{(\omega(T_i)\lambda_0(T_i))'}{\lambda_0(T_i)}\left(\frac{\Delta_i}{Y(T_i)/n} - \frac{\Delta_i}{S_T(T_i)}\right)\mathbb{1}_{\{T_i> t_\varepsilon\}}, \quad (31)\end{aligned}$$

where $\varepsilon > 0$ and $t_\varepsilon > 0$, and t_ϵ is the infimum over all $t > 0$ such that

$$\int_t^\infty \int_t^\infty \frac{|K^\star(t,s)|}{\lambda_0(t)\lambda_0(s)S_T(t)S_T(s)} S_C(t)S_C(s)f_X(t)f_X(s)dtds \leq \varepsilon.$$

Notice that such a t_ϵ is well-defined by Lemma 9 and Condition d.3). For the term in Equation (30), observe that

$$\left(\sup_{\omega\in B_1(\mathcal{H}^{(p)})} \frac{1}{n}\sum_{i=1}^n \frac{(\omega(T_i)\lambda_0(T_i))'}{\lambda_0(T_i)}\left(\frac{\Delta_i}{Y(T_i)/n} - \frac{\Delta_i}{S_T(T_i)}\right)\mathbb{1}_{\{T_i\leq t_\varepsilon\}}\right)^2 \tag{32}$$
$$\leq \sup_{t\leq t_\varepsilon}\left(\frac{1}{Y(t)/n} - \frac{1}{S_T(t)}\right)^2 \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \Delta_i\Delta_j \frac{K^\star(T_i,T_j)}{\lambda_0(T_i)\lambda_0(T_j)}\mathbb{1}_{\{T_i\leq t_\varepsilon\}}\mathbb{1}_{\{T_j\leq t_\varepsilon\}}$$
$$= o_p(1) \tag{33}$$

where the last line holds since $\sup_{t\leq t_\epsilon}\left|\frac{1}{Y(t)/n} - \frac{1}{S_T(t)}\right| = o_p(1)$ a.s., by an application of Glivenko-Cantelli, and since the double sum converges to

$$\mathbb{E}\left(\Delta_1\Delta_2\frac{K^\star(T_1,T_2)}{\lambda_0(T_1)\lambda_0(T_2)}\mathbb{1}_{\{T_1\leq t_\varepsilon\}}\mathbb{1}_{\{T_2\leq t_\varepsilon\}}\right),$$

which is finite by Lemma 9 and Condition d.3).

Finally, we prove that the term in Equation (31) is $o_p(1)$. Define $R(t) = \left|\frac{S_T(t)}{Y(t)/n} - 1\right|$. Gill (1983) proved that $\sup_{t\leq\tau_n} R(t) = O_p(1)$ where $\tau_n = \max\{T_1,\ldots,T_n\}$. By using this result, the term in Equation (31) satisfies

$$\left(\sup_{\omega\in B_1(\mathcal{H}^{(p)})} \frac{1}{n}\sum_{i=1}^n \frac{(\omega(T_i)\lambda_0(T_i))'}{\lambda_0(T_i)}\left(\frac{\Delta_i}{Y(T_i)/n} - \frac{\Delta_i}{S_T(T_i)}\right)\mathbb{1}_{\{T_i> t_\varepsilon\}}\right)^2$$
$$\leq \frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \frac{\Delta_i\Delta_j|K^\star(T_i,T_j)|}{\lambda_0(T_i)\lambda_0(T_j)S_T(T_i)S_T(T_j)}R(T_i)R(T_j)\mathbb{1}_{\{T_i> t_\varepsilon\}}\mathbb{1}_{\{T_j> t_\varepsilon\}}$$
$$= O_p(1)\frac{1}{n^2}\sum_{i=1}^n\sum_{j=1}^n \frac{\Delta_i\Delta_j|K^\star(T_i,T_j)|}{\lambda_0(T_i)\lambda_0(T_j)S_T(T_i)S_T(T_j)}\mathbb{1}_{\{T_i> t_\varepsilon\}}\mathbb{1}_{\{T_j> t_\varepsilon\}}$$
$$= O_p(1)\int_{t_\varepsilon}^\infty\int_{t_\varepsilon}^\infty \frac{|K^\star(t,s)|}{\lambda_0(t)\lambda_0(s)S_T(t)S_T(s)}S_C(t)S_C(s)f_X(t)f_X(s)dtds$$
$$= O_p(1)\varepsilon,$$

where in the second line we used that $\sup_{t\leq\tau_n} R(t) = O_p(1)$, and in the fourth line we used the law of large numbers, and the definition of t_ε. Since ε is arbitrary, we conclude that equation (31) tends to 0 in probability.

A.4.2. Proof of Theorem 7

Survival Stein operator (c=s): We proceed by contradiction. Assume that $f_X \neq f_0$ but c-KSD$(f_X\|f_0) = \sup_{\omega\in B_1(\mathcal{H}^{(s)})}\mathbb{E}_X((\mathcal{T}_0^{(s)}\omega)(T,\Delta)) = 0$. Recall that

$$\mathbb{E}_X((\mathcal{T}_0^{(s)}\omega)(T,\Delta))$$
$$= \mathbb{E}_X((\mathcal{T}_0\omega)(T,\Delta))$$
$$= \mathbb{E}_X\left[\Delta\omega'(T) + \Delta\omega(T)\frac{f_0'(T)}{f_0(T)} - \Delta\omega(T)\lambda_C(T)\right] + \omega(0)f_0(0).$$

Similarly, define

$$(\mathcal{T}_X\omega)(x,\delta) = \delta\omega'(x) + \delta\omega(x)\frac{f_X'(x)}{f_X(x)} - \delta\omega(x)\lambda_C(x) + \omega(0)f_X(0),$$

and notice that $\mathbb{E}_X((\mathcal{T}_X\omega)(T,\Delta)) = 0$ by the Stein's identity. Then

$$\begin{aligned}\mathbb{E}_X\left((\mathcal{T}_0^{(s)}\omega)(T,\Delta)\right) &= \mathbb{E}_X\left((\mathcal{T}_0\omega)(T,\Delta)\right)\\ &= \mathbb{E}_X\left((\mathcal{T}_0\omega)(T,\Delta) - (\mathcal{T}_X\omega)(T,\Delta)\right)\\ &= \mathbb{E}_X\left(\Delta\omega(T)\left(\frac{f_0'(T)}{f_0(T)} - \frac{f_X'(T)}{f_X(T)}\right) + \omega(0)(f_0(0) - f_X(0))\right)\\ &= \mathbb{E}_X\left(\Delta\omega(T)\left(\log\frac{f_0(T)}{f_X(T)}\right)'\right) + \omega(0)(f_0(0) - f_X(0)),\end{aligned}$$

and thus, we have

$$\begin{aligned}0 = \text{s-KSD}(f_X\|f_0) &= \sup_{\omega\in B_1(\mathcal{H}^{(s)})}\mathbb{E}_X((\mathcal{T}_0^{(s)}\omega)(T,\Delta))\\ &= \sup_{\omega\in B_1(\mathcal{H}^{(s)})}\mathbb{E}_X\left(\Delta\omega(T)\left(\log\frac{f_0(T)}{f_X(T)}\right)'\right) + \omega(0)(f_0(0) - f_X(0))\\ &= \sup_{\omega\in B_1(\mathcal{H}^{(s)})}\left\langle\omega, \int_0^\infty K(x,\cdot)d\nu(x)\right\rangle = \left\|\int_0^\infty K(x,\cdot)d\nu(x)\right\|_{\mathcal{H}^{(s)}},\end{aligned}$$

where $d\nu(x) = \left(\log\frac{f_0(x)}{f_X(x)}\right)' S_C(x)f_X(x)dx + (f_0(x) - f_X(x))\delta_0(x)$, and where we identify $\int_0^\infty K(x,\cdot)d\nu(x)$ as the mean kernel embedding of the measure ν. We shall assume that the above embedding is well-defined, otherwise we have s-KSD$(f_X\|f_0) \neq 0$. Since the kernel is c_0-universal, the previous set of equations implies ν is the zero measure, which implies that $f_0(0) = f_X(0)$, and

$$\left(\log\frac{f_0(x)}{f_X(x)}\right)' = 0, \tag{34}$$

as long as $f_X(x) > 0$ implies $S_C(x)f_X(x) > 0$ (which does, since we assume $S_C(x) = 0$ implies $S_X(x) = \int_x^\infty f_X(x)dx = 0$). Equation (34) yields $f_0 \propto f_X$ and $f_X = f_0$ since both, f_0 and f_X, are probability density functions. This finalizes our proof.

Martingale Stein operator (c=m): Define

$$(\mathcal{T}_X^{(m)}\omega)(x,\delta) = \omega'(x)\frac{\delta}{\lambda_X(x)} - (\omega(x) - \omega(0)),$$

and notice that $\mathbb{E}_X((\mathcal{T}_X^{(m)}\omega)(T,\Delta)) = 0$ follows from the martingale identity. Observe that

$$\begin{aligned}\text{m-KSD}(f_X\|f_0) &= \sup_{\omega\in B_1(\mathcal{H}^{(m)})}\mathbb{E}_X((\mathcal{T}_0^{(m)}\omega)(T,\Delta))\\ &= \sup_{\omega\in B_1(\mathcal{H}^{(m)})}\mathbb{E}_X((\mathcal{T}_0^{(m)}\omega)(T,\Delta)) - \mathbb{E}_X((\mathcal{T}_X^{(m)}\omega)(T,\Delta))\\ &= \sup_{\omega\in B_1(\mathcal{H}^{(m)})}\mathbb{E}_X\left(\omega'(T)\Delta\left(\frac{1}{\lambda_0(T)} - \frac{1}{\lambda_X(T)}\right)\right)\\ &= \sup_{\omega\in B_1(\mathcal{H}^{(m)})}\int_0^\infty \omega'(x)\left(\frac{1}{\lambda_0(x)} - \frac{1}{\lambda_X(x)}\right)f_X(x)S_C(x)dx.\end{aligned}$$

Denote $\alpha(x) = \left(\frac{1}{\lambda_0(x)} - \frac{1}{\lambda_X(x)}\right)f_X(x)S_C(x)$, and, as usual, $K^\star(x,y) = \frac{\partial^2}{\partial x\partial y}K(x,y)$. Then,

$$\text{m-KSD}^2(f_X\|f_0) = \int_0^\infty\int_0^\infty \alpha(x)K^\star(x,y)\alpha(y)dxdy.$$

Since $K^\star$ is c_0-universal by Condition a), the previous term is equal to 0 if and only if $\alpha(x) = 0$ for all $x > 0$. Now, $\alpha(x) = 0$ if and only if $\frac{1}{\lambda_0(x)} - \frac{1}{\lambda_X(x)} = 0$, which holds if and only if $f_0(x) = f_X(x)$ for all $x > 0$.

A.4.3. Proof of Theorem 8

Deterministic kernels ($c \in \{s, m\}$)**:** For $c \in \{s, m\}$ which are associated to a deterministic kernel function $h^{(c)}((T, \Delta), (T', \Delta'))$, the result follows from the classical theory of V-statistics since $h^{(c)}$ are degenerate kernels, and under the following moment conditions:

i) $\mathbb{E}_0(|h^{(c)}((T, \Delta), (T, \Delta))|) < \infty$, and

ii) $\mathbb{E}_0(h^{(c)}((T, \Delta), (T', \Delta'))^2) < \infty$,

which are satisfied due to Lemma 10.

Random kernel ($c \in \{p\}$)**:** Observe that

$$\begin{aligned}\sqrt{n}\widehat{\text{c-KSD}}(f_X \| f_0) &= \sup_{\omega \in B_1(\mathcal{H}^{(p)})} \frac{1}{\sqrt{n}} \sum_{i=1}^n \frac{(\omega(T_i)\lambda_0(T_i))'}{\lambda_0(T_i)} \frac{\Delta_i}{Y(T_i)/n} \\ &= \sup_{\omega \in B_1(\mathcal{H}^{(p)})} \frac{1}{\sqrt{n}} \int_0^{\tau_n} \frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)} \frac{1}{Y(x)/n} dN(x),\end{aligned}$$

where $dN(x) = \sum_{i=1}^n \Delta_i \delta_{T_i}(x)$. By hypothesis, $\int_0^\infty (\omega(x)\lambda_0(x))' dx = 0$ for all $\omega \in \mathcal{H}^{(p)}$, then

$$\begin{aligned}\sqrt{n}\widehat{\text{c-KSD}}(f_X \| f_0) &= \sup_{\omega \in B_1(\mathcal{H}^{(p)})} \frac{1}{\sqrt{n}} \int_0^{\tau_n} \frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)} \frac{1}{Y(x)/n} dN(x) - \sqrt{n} \int_0^\infty (\omega(x)\lambda_0(x))' dx \\ &= \sup_{\omega \in B_1(\mathcal{H}^{(p)})} \frac{1}{\sqrt{n}} \int_0^{\tau_n} \frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)} \frac{1}{Y(x)/n} dM(x) - \sqrt{n} \int_{\tau_n}^\infty (\omega(x)\lambda_0(x))' dx\end{aligned}$$

where $dM(x) = dN(x) - Y(x)\lambda_0(x)dx$. Therefore we conclude that $\sqrt{n}\widehat{\text{c-KSD}}(f_X \| f_0) \in [a - b, a + b]$, where

$$\begin{aligned}a &= \sup_{\omega \in B_1(\mathcal{H}^{(p)})} \frac{1}{\sqrt{n}} \int_0^{\tau_n} \frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)} \frac{1}{Y(x)/n} dM(x), \text{ and} \\ b &= \sup_{\omega \in B_1(\mathcal{H}^{(p)})} \sqrt{n} \int_{\tau_n}^\infty (\omega(x)\lambda_0(x))' dx\end{aligned}$$

We will prove that $b = o_p(1)$. Let $K^\star(x, y) = \left(\frac{\partial^2}{\partial x \partial y} \lambda_0(x)\lambda_0(y)K(x, y)\right)$, then

$$\begin{aligned}\left(\sup_{\omega \in B_1(\mathcal{H}^{(p)})} \sqrt{n} \int_{\tau_n}^\infty (\omega(x)\lambda_0(x))' dx\right)^2 &= n \int_{\tau_n}^\infty \int_{\tau_n}^\infty \frac{K^\star(x, y)}{f_T(x) f_T(y)} f_T(x) f_T(y) dx dy \\ &\leq n S_T(\tau_n)^{1/2} \left(\int_{\tau_n}^\infty \left(\int_{\tau_n}^\infty \frac{K^\star(x, y)}{f_T(x) f_T(y)} f_T(x) dx\right)^2 f_T(y) dy\right)^{1/2} \\ &\leq n S_T(\tau_n) \left(\int_{\tau_n}^\infty \int_{\tau_n}^\infty \frac{K^\star(x, y)^2}{f_T(x)^2 f_T(y)^2} f_T(x) f_T(y) dx dy\right)^{1/2},\end{aligned}$$

where the two inequalities above follow from the Cauchy-Schwarz inequality, by the fact that $nS_T(\tau_n) = O_p(1)$ (Yang, 1994), and the previous double integral converges to 0 by Condition c.3), since $\tau_n = \max\{T_1, \ldots, T_n\} \to \infty$. From the previous result, we deduce

$$\sqrt{n}\widehat{\text{c-KSD}}(f_X \| f_0) = \sup_{\omega \in B_1(\mathcal{H}^{(p)})} \frac{1}{\sqrt{n}} \int_0^{\tau_n} \frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)} \frac{1}{Y(x)/n} dM(x) + o_p(1).$$

The previous step is important in our analysis as it allows us to write $\sqrt{n}\widehat{\text{c-KSD}}(f_X\|f_0)$ in terms of $M(x)$. Our next step is to prove that we can replace the term $Y(x)/n$, in the previous equation, by $S_T(x)$. Observe

$$\begin{aligned}
&\sqrt{n}\widehat{\text{c-KSD}}(f_X\|f_0)\\
&= \sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{\sqrt{n}}\int_0^{\tau_n}\frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)}\left(\frac{1}{Y(x)/n}-\frac{1}{S_T(x)}+\frac{1}{S_T(x)}\right)dM(x)+o_p(1)\\
&= \sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{\sqrt{n}}\int_0^{\tau_n}\frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)}\frac{1}{S_T(x)}dM(x)\\
&\quad\pm \sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{\sqrt{n}}\int_0^{\tau_n}\frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)}\left(\frac{1}{Y(x)/n}-\frac{1}{S_T(x)}\right)dM(x)+o_p(1).
\end{aligned}$$

The $\pm$ notation above denotes lower, given by $-$, and upper, given by $+$, bounds for $\sqrt{n}\widehat{\text{c-KSD}}(f_X\|f_0)$. Finally, by taking square, the result is deduced by proving

$$\sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{\sqrt{n}}\int_0^{\tau_n}\frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)}\left(\frac{1}{Y(x)/n}-\frac{1}{S_T(x)}\right)dM(x)=o_p(1),$$

and

$$\sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{\sqrt{n}}\int_0^{\tau_n}\frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)}\frac{1}{S_T(x)}dM(x)=O_p(1).$$

The second equation won't be verified as, at the end of this proof, we will show that such a quantity converges in distribution to some random variable, thus it will be bounded in probability. For the first equation, notice that

$$\begin{aligned}
&\left(\sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{\sqrt{n}}\int_0^{\tau_n}\frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)}\left(\frac{1}{Y(x)/n}-\frac{1}{S_T(x)}\right)dM(x)\right)^2\\
&=\frac{1}{n}\int_0^{\tau_n}\int_0^{\tau_n}\frac{K^\star(x,y)}{\lambda_0(x)\lambda_0(y)}\left(\frac{1}{Y(x)/n}-\frac{1}{S_T(x)}\right)\left(\frac{1}{Y(y)/n}-\frac{1}{S_T(y)}\right)dM(x)dM(y),
\end{aligned}$$

is a double integral with respect to the $M(x)$. Then, by Theorem 17 of Fernandez & Rivera (2019), it is enough to verify

$$\frac{1}{n}\int_0^{\tau_n}\frac{K^\star(x,x)}{\lambda_0(x)^2}\left(\frac{1}{Y(x)/n}-\frac{1}{S_T(x)}\right)^2Y(x)\lambda_0(x)dx=o_p(1).$$

Observe that

$$\begin{aligned}
\frac{1}{n}\int_0^{\tau_n}\frac{K^\star(x,x)}{\lambda_0(x)^2}\left(\frac{1}{Y(x)/n}-\frac{1}{S_T(x)}\right)^2Y(x)\lambda_0(x)dx&=\int_0^{\tau_n}\frac{K^\star(x,x)}{\lambda_0(x)^2}\left(1-\frac{Y(x)/n}{S_T(x)}\right)^2\frac{1}{Y(x)/n}\lambda_0(x)dx\\
&=O_p(1)\int_0^{\tau}\frac{K^\star(x,x)}{\lambda_0(x)^2}\left(1-\frac{Y(x)/n}{S_T(x)}\right)^2\frac{1}{S_T(x)}\lambda_0(x)dx\\
&=o_p(1),
\end{aligned}$$

where the second equality follows from $n/Y(x)=O_p(1)1/S_T(x)$ uniformly for all $x\leq\tau_n$ (Gill, 1983), and the last equality is due to dominated convergence in sets of probability as high as desired, as $\left(1-\frac{Y(x)/n}{S_T(x)}\right)\to 0$ for all $x<\infty$ from the Glivenko Cantelli Theorem, and

$$\frac{K^\star(x,x)}{\lambda_0(x)^2}\left(1-\frac{Y(x)/n}{S_T(x)}\right)^2\frac{1}{S_T(x)}\lambda_0(x)=O_p(1)\frac{K^\star(x,x)}{f_0(x)^2S_C(x)}f_0(x),$$

which is integrable by Condition c.3).

Putting everything together, we have shown that

$$\begin{aligned}\sqrt{n}\widehat{\text{c-KSD}}^2(f_X\|f_0) &= \left(\sup_{\omega\in B_1(\mathcal{H}^{(p)})}\frac{1}{\sqrt{n}}\int_0^{\tau_n}\frac{(\omega(x)\lambda_0(x))'}{\lambda_0(x)}\frac{1}{S_T(x)}dM(x)\right)^2+o_p(1)\\ &= \frac{1}{n}\int_0^{\tau_n}\int_0^{\tau_n}\frac{K^\star(x,y)}{f_0(x)f_0(y)S_C(x)S_C(y)}dM(x)dM(y)+o_p(1)\\ &= \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n\int_0^{X_i}\int_0^{X_j}\frac{K^\star(x,y)}{f_0(x)f_0(y)S_C(x)S_C(y)}dM_j(x)dM_i(y)+o_p(1)\\ &= \frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n J((T_i,\Delta_i),(T_j,\Delta_j))+o_p(1),\end{aligned}$$

where $M_i(x)=N_i(x)-\int_0^x \mathbb{1}_{\{T_i\geq y\}}\lambda_0(y)dy=\Delta_i\mathbb{1}_{\{T_i\leq x\}}-\int_0^x\mathbb{1}_{\{T_i\geq y\}}\lambda_0(y)dy$. Notice that the process $M_i(x)$ only depends on the i-th observation (T_i,Δ_i). Notice that the previous expression is approximately a V-statistic with kernel given by $J((T_i,\Delta_i),(T_j,\Delta_j))=\int_0^{T_i}\int_0^{T_j}\frac{K^\star(x,y)}{f_0(x)f_0(y)S_C(x)S_C(y)}dM_j(x)dM_i(y)$. By proposition 23 of Fernandez & Rivera (2019), we have that $\mathbb{E}(J((T_i,\Delta_i),(T_j,\Delta_j))|T_i,\Delta_i)=0$, thus J is a degenerate V-statistic kernel.

By the classical theory of V-statistics,

$$\frac{1}{n}\sum_{i=1}^n\sum_{j=1}^n J((T_i,\Delta_i),(T_j,\Delta_j))\overset{\mathcal{D}}{\to} r_p+\mathcal{Y}_p,$$

where r_p is a constant and $\mathcal{Y}_p$ is a (potentially) infinite sum of independent χ^2 random variables, as long as the following moment conditions are satisfied:

$$\text{i) } \mathbb{E}_0(|J((T_1,\Delta_1),(T_1,\Delta_1))|)<\infty, \quad \text{and} \quad \text{ii) } \mathbb{E}_0(J((T_1,\Delta_1),(T_2,\Delta_2))^2)<\infty.$$

Again, by Proposition 23 of Fernandez & Rivera (2019), checking those moment conditions is equivalent to verify:

$$\text{i) } \mathbb{E}_0\left[\frac{K^\star(T,T)\Delta}{(f_0(T)S_C(T))^2}\right]<\infty \quad \text{and} \quad \text{ii) } \mathbb{E}_0\left[\frac{K^\star(T,T')^2\Delta\Delta'}{(f_0(T)f_0(T')S_C(T)S_C(T'))^2}\right]<\infty,$$

which are exactly the conditions assumed in Condition c.3).

B. Known Identities

In Section 3.2, to derive the martingale Stein operator, we use the following identity

$$\mathbb{E}_0\left[\Delta\phi(T)-\int_0^T\phi(t)\lambda_0(t)dt\right]=0,$$

which holds under the null hypothesis, where λ_0 is the hazard function under the null.

Let $N_i(x)$ and $Y_i(x)$ be the individual counting and risk processes, defined by by $N_i(x)=\Delta_i\mathbb{1}_{\{T_i\leq x\}}$ and $Y_i(x)=\mathbb{1}_{\{T_i\geq x\}}$, respectively. Then, the individual zero-mean martingale for the i-th individual corresponds to $M_i(x)=N_i(x)-\int_0^x Y_i(y)\lambda_0(y)dy$, where $\mathbb{E}_0(M_i(x))=0$ for all x.

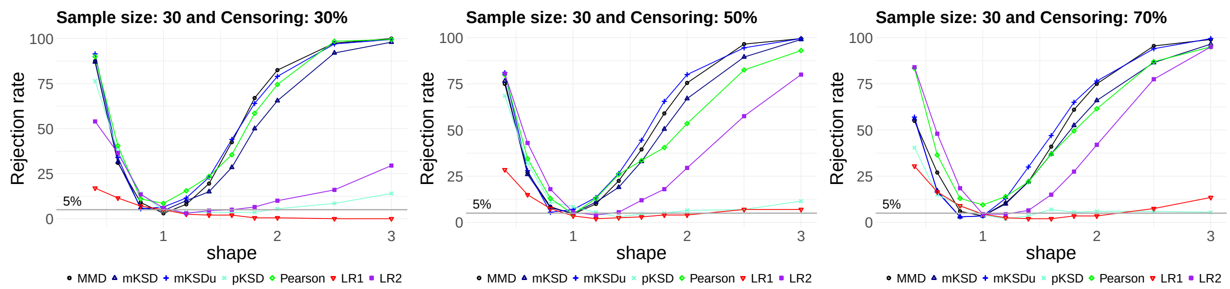
Additionally, let $\phi:\mathbb{R}_+\to\mathbb{R}$ such that $\mathbb{E}_0\left|\int_0^x\phi(y)dM_i(y)\right|<\infty$ for all x, then $\int_0^x\phi(y)dM_i(y)$ is a zero-mean $(\mathcal{F}_x)$-martingale (See Chapter 2 of (Aalen et al., 2008)). The, taking expectation, we have

$$\mathbb{E}_0\left[\int_0^\infty\phi(x)dM_i(x)\right]=\mathbb{E}_0\left[\int_0^\infty\phi(x)(dN_i(x)-Y_i(x)\lambda_0(x)dx)\right]=\mathbb{E}_0\left[\Delta\phi(T)-\int_0^T\phi(x)\lambda_0(x)dx\right]=0.$$

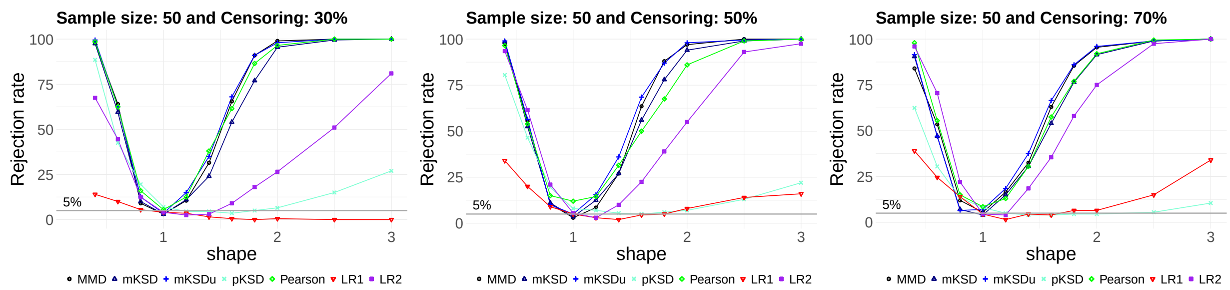
C. Additional Experiments

C.1. Weibull experiments: small deviations from the null

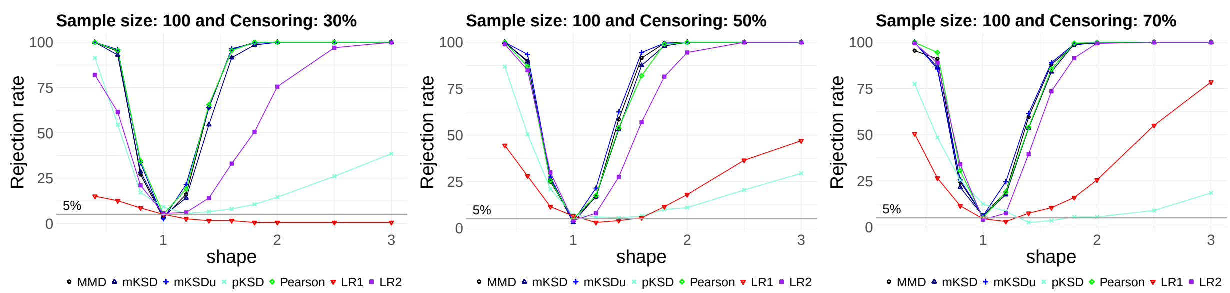
SAMPLE SIZE: 30, AND CENSORING PERCENTAGES OF 30%, 50% AND 70%



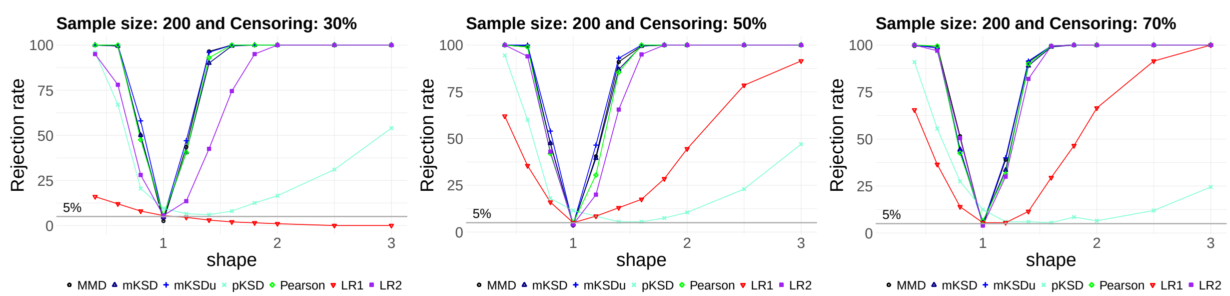
SAMPLE SIZE: 50, AND CENSORING PERCENTAGES OF 30%, 50% AND 70%



SAMPLE SIZE: 100, AND CENSORING PERCENTAGES OF 30%, 50% AND 70%

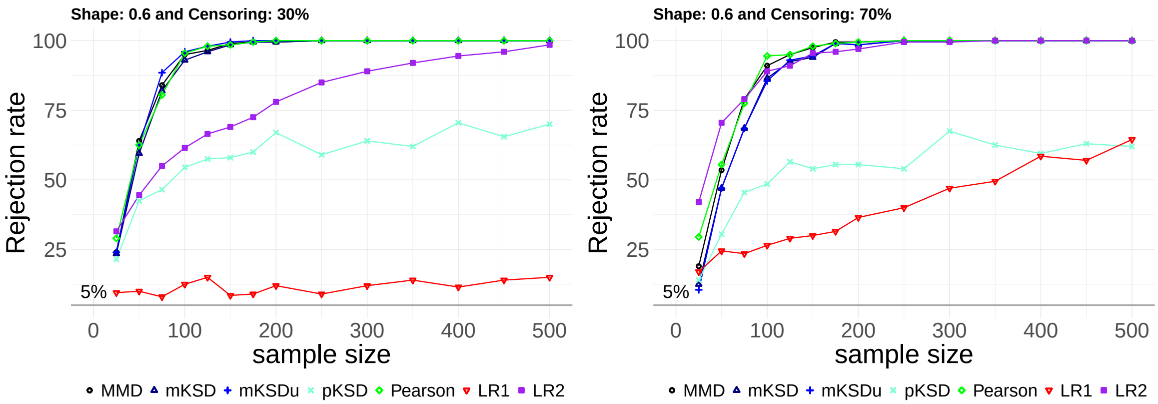


SAMPLE SIZE: 200, AND CENSORING PERCENTAGES OF 30%, 50% AND 70%

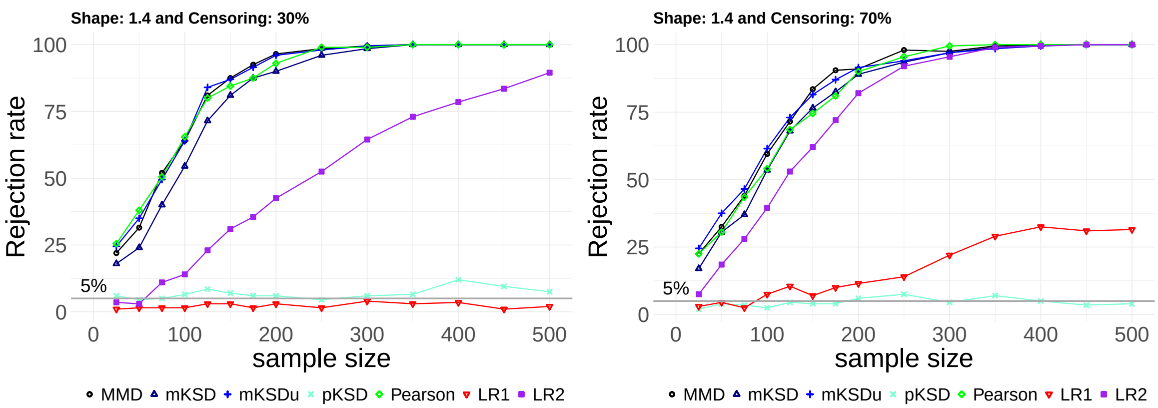


C.2. Weibull experiments: increasing sample size

SHAPE: 0.6 AND CENSORING PERCENTAGES OF 30% AND 70%

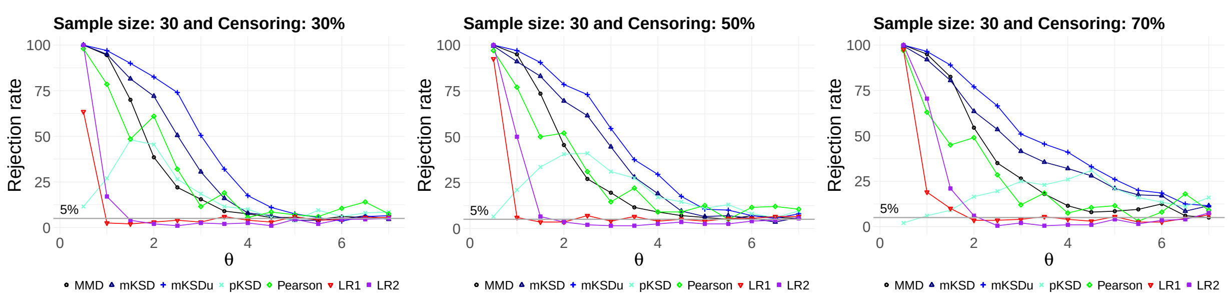


SHAPE: 1.4 AND CENSORING PERCENTAGES OF 30% AND 70%

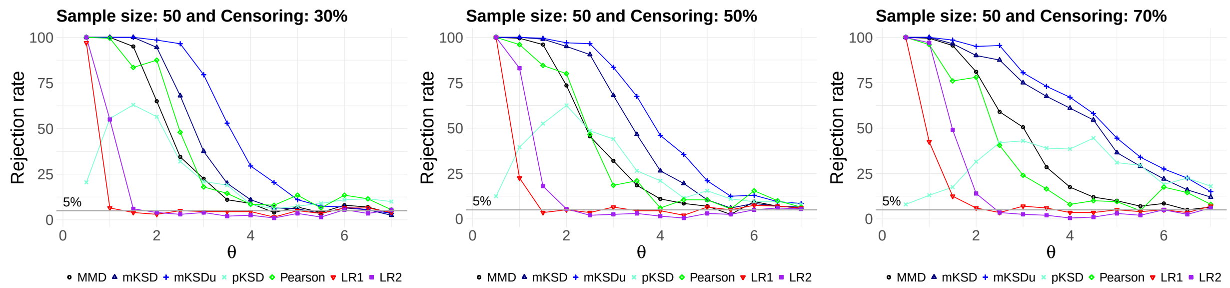


C.3. Periodic experiments: small deviations from the null

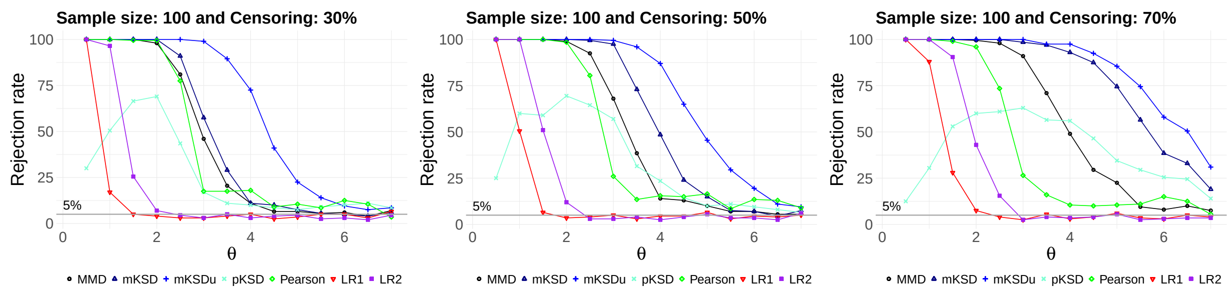
SAMPLE SIZE: 30, AND CENSORING PERCENTAGES OF 30%, 50% AND 70%



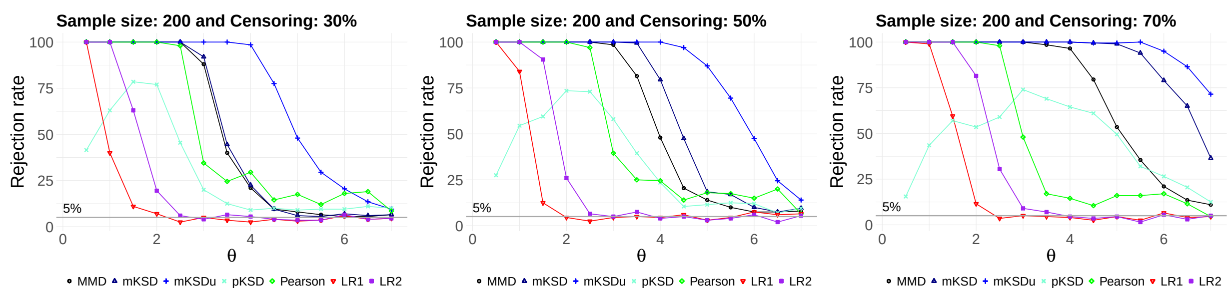
SAMPLE SIZE: 50, AND CENSORING PERCENTAGES OF 30%, 50% AND 70%



SAMPLE SIZE: 100, AND CENSORING PERCENTAGES OF 30%, 50% AND 70%

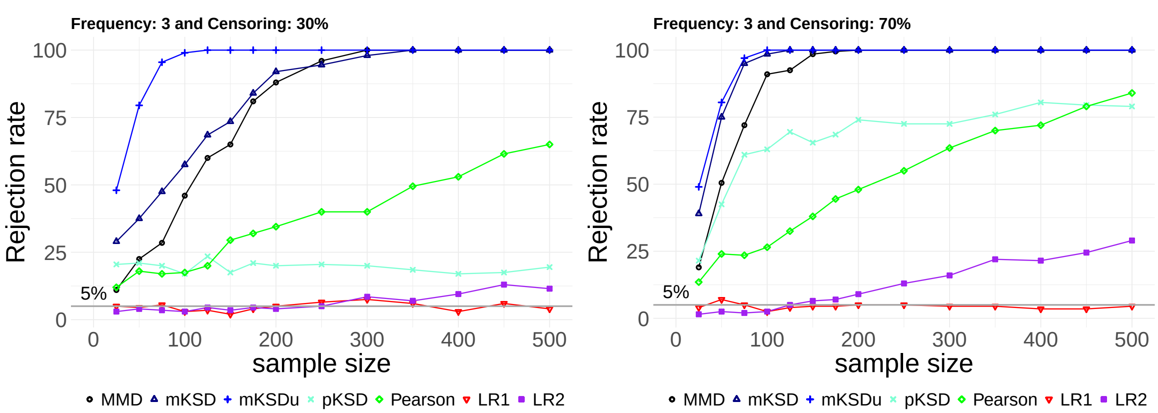


SAMPLE SIZE: 200, AND CENSORING PERCENTAGES OF 30%, 50% AND 70%



C.4. Periodic experiments: increasing sample size

FREQUENCY: 3 AND CENSORING PERCENTAGES OF 30% AND 70%



FREQUENCY: 6 AND CENSORING PERCENTAGES OF 30% AND 70%